

# ChemComm

Chemical Communications

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: A. Kumar and N. Sabouri, *Chem. Commun.*, 2026, DOI: 10.1039/D6CC01453D.



This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about Accepted Manuscripts in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.

## COMMUNICATION

Direct visualization and functional characterization of *Gar2<sup>Nucleolin</sup>*-G-quadruplex interactionsReceived 00th January 20xx,  
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Ajay Kumar,<sup>a</sup> and Nasim Sabouri,<sup>a,b,c</sup>

We demonstrate that *Gar2<sup>Nucleolin</sup>*, a conserved nucleolar protein required for ribosome biogenesis in *Schizosaccharomyces pombe*, interacts with both DNA and RNA G-quadruplexes (G4s), and destabilize DNA G4 structures. Using biochemical assays and, G4-immuno-electron microscopy, we directly visualize *Gar2*'s interaction with G4s at high resolution. Our findings establish *Gar2* as a novel regulator of G4 dynamics, expanding the role of nucleolar proteins in shaping genome architecture.

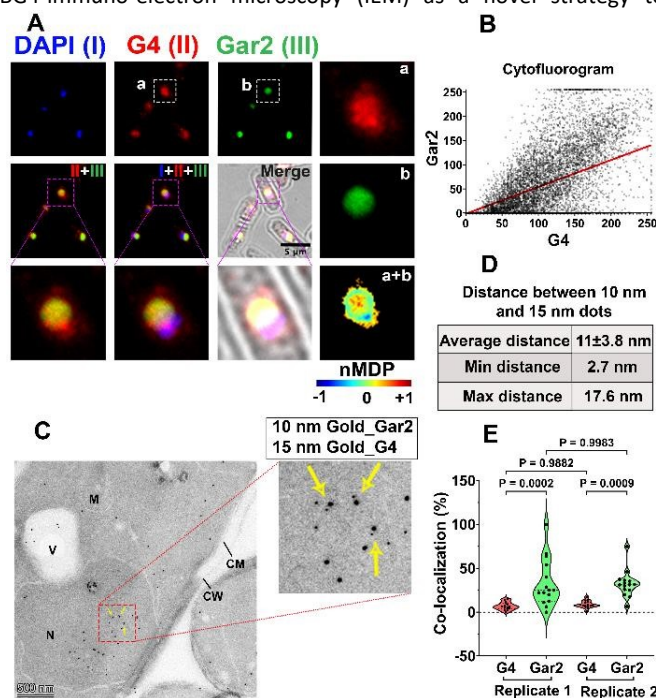
Guanine (G)-rich sequences in nucleic acids can adopt stable non-canonical four-stranded structures known as G-quadruplexes (G4s). G4s are composed of stacked G-quartets stabilized by Hoogsteen hydrogen bonding and  $\pi$ - $\pi$  interactions<sup>1</sup>. Monovalent cations particularly  $K^+$ , further enhance G4 stability by coordinating with G bases<sup>2</sup>. Sequences with potential to form G4s are evolutionary conserved and enriched in regulatory genomic regions such as telomeres, promoters, ribosomal DNA (rDNA) and untranslated regions (UTRs), where they influence processes including DNA replication, transcription, translation, and epigenetic regulation<sup>3</sup>. Dysregulation of G4s is associated with various pathologies, including neurodegenerative diseases (e.g., ALS)<sup>4</sup>, cancers<sup>5</sup>, cardiovascular disorders<sup>6</sup> and viral infections as G4s are also present in key viral genes<sup>7</sup>.

G4 dynamics are regulated by environmental factors such as  $K^+$ , and  $Na^+$  ions<sup>8</sup>, chromatin structure<sup>9</sup>, and protein interactions<sup>10</sup>. G4-binding proteins (G4BPs) are broadly categorized into those that bind DNA and RNA G4s, and can function as recruiters or effectors. Recruiters, such as SP1, binds G4 sites and regulate gene expression by facilitating binding of transcription factors at G4 sites<sup>11</sup>. Effectors either stabilize (e.g., Nucleolin (NCL))<sup>12</sup> or destabilize (e.g. hnRNP A1) G4 structures<sup>13</sup>, often through conserved domains such as arginine-glycine repeats (RGG motifs), and RNA recognition motifs (RRM)<sup>14</sup>.

While numerous G4BPs with these motifs have been characterized in mammalian systems, their counterparts in unicellular eukaryotes remain largely unexplored<sup>15</sup>, limiting our understanding of conserved principles of G4 regulation. To address this gap, we investigated potential G4BPs in the fission yeast *Schizosaccharomyces pombe* and identified *Gar2* as a candidate, a conserved nucleolar protein essential for ribosome biogenesis and homologous to human NCL, a well-characterized G4BP with four RRM and an RGG<sup>16, 17</sup>. Structural analysis of NCL in complex with

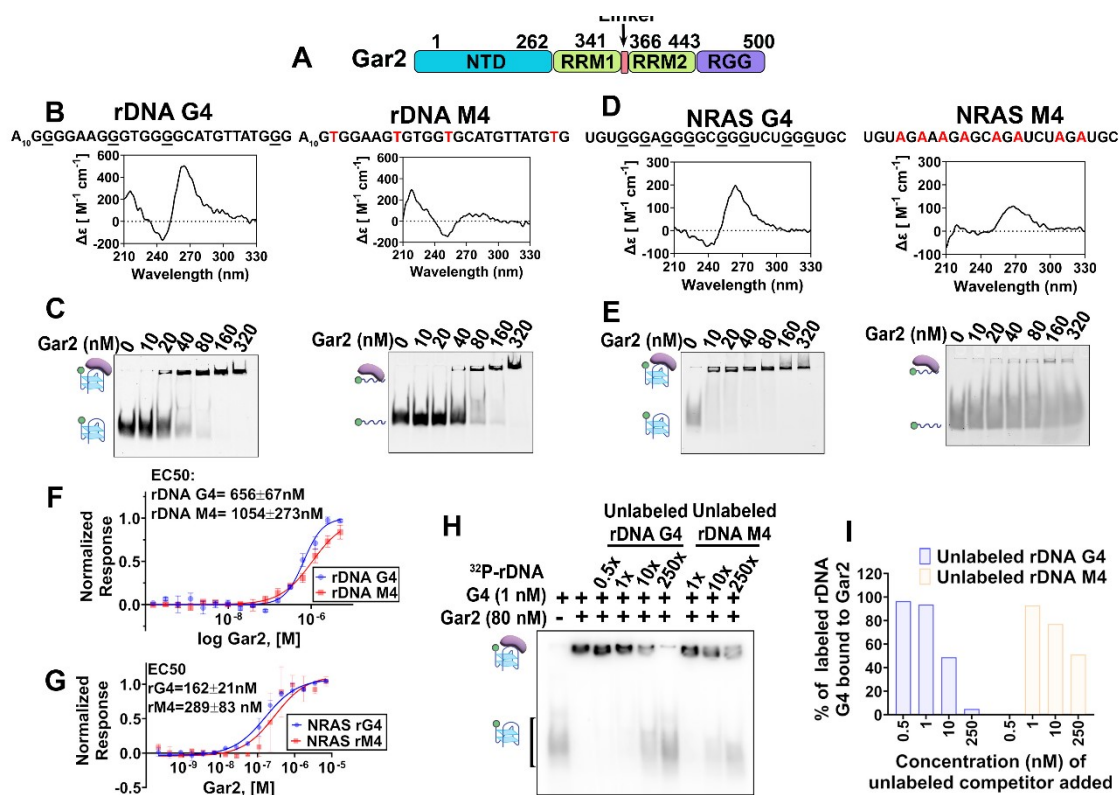
the c-MYC G4 demonstrated that all four RRM cooperate in high-affinity G4 recognition<sup>17</sup>. Notably, NCL is frequently overexpressed in cancer, where it has been linked to G4-mediated regulation of oncogene expression<sup>18</sup>.

We systematically dissected the G4-binding properties of full-length *Gar2* and its domain-deletion variants. Importantly, we introduced BG4-immuno-electron microscopy (IEM) as a novel strategy to



**Figure 1. Visualization of G4s and *Gar2*.** (A) Immunofluorescence microscopy demonstrates colocalization of *Gar2* with G4s. Top row: single-label images of nucleus (DAPI, blue), G4s (red), and *Gar2* (green). Middle row: merged images of *Gar2*/G4s, nucleus/*Gar2*/G4s, and brightfield overlay. Bottom row: magnified regions (magenta boxes); colormap (nMDP) depicts color bar for colocalization intensity. (B) Scatter plot of G4 and *Gar2* pixel intensities from cells shown in (A). (C) Immunoelectron microscopy demonstrates colocalization of *Gar2* (10 nm gold) and G4s (15 nm gold); yellow arrows mark proximity, inset shows amplified view of a region. Abbreviations: N=nucleus, M=mitochondria, V=vacuole, CM=cell membrane, CW=cell wall. (D) Table: average distance from colocalization events (N=50). (E) Violin plot showing distribution of colocalization percentages of *Gar2* and G4s across two biological replicates. Dots = individual cells. Statistics: one-way ANOVA, Tukey's multiple comparisons test; P-values indicated.

<sup>a</sup> Department of Medical Biochemistry and Biophysics, Umeå 90187, Sweden<sup>b</sup> Science for Life Laboratory, Umeå University, Umeå, Sweden<sup>c</sup> Corresponding E-mail: nasim.sabouri@umu.se

View Article Online  
DOI: 10.1039/D6CC01453D

**Figure 2. Gar2 binds G4 structures.** (A) Schematic of Gar2 domains (B) CD spectra of rDNA G4 and rDNA M4 oligonucleotides; mutated nucleotides underlined and in red. (C) EMSA shows Gar2 binding to Cy5-labeled rDNA G4 and rDNA M4. (D) Circular dichroism spectra of NRAS G4 RNA and NRAS M4 RNA oligonucleotides; sequences as in (B). (E) EMSA shows Gar2 binding to Cy5-labeled NRAS G4 and NRAS M4 RNA. (F, G) Quantification of Gar2 affinity towards rDNA, rDNA M4, NRAS G4 (rG4) and NRAS M4 (rM4) oligonucleotides by MST (two replicates, EC50 by Hill equation). (H) Competitive EMSA shows Gar2 prefers rDNA G4 over rDNA M4. (I) Fraction of bound probe quantified from EMSA gels in (H), plotted as percent bound vs competitor concentration.

directly visualize and spatially resolve G4–Gar2 interactions at high resolution. Our data demonstrate that Gar2 binds both DNA and RNA G4s. Further we show that Gar2 destabilizes G4 DNA structures indicating a previously uncharacterized role for Gar2 in G4 biology and suggesting an expanded function for nucleolar proteins in regulating genome architecture in *S. pombe*. Unlike its homolog NCL, which stabilizes G4s, Gar2 promotes G4 destabilization, highlighting species-specific and context-dependent mechanisms of G4 regulation.

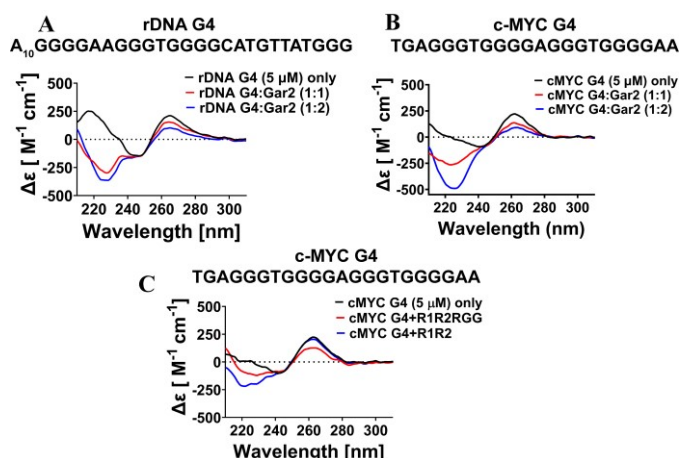
To investigate whether Gar2 interacts with G4s *in vivo*, we examined its subcellular localization with G4 structures in *S. pombe*. To visualize Gar2 *in vivo*, the endogenous *gar2* gene was C-terminally tagged with a 3×HA (hemagglutinin) epitope (Gar2-HA) by homologous recombination (for details, see Supporting Information). Using Gar2-HA cells (Fig. S1), we performed immunofluorescence microscopy combined with BG4 antibody staining, which recognizes both DNA and RNA G4s (Fig. S2)<sup>19</sup>. While G4 formation in *S. pombe* has been detected by ChIP, G4 visualization by microscopy has not been reported<sup>20–22</sup>. BG4 staining was observed in both the nucleus and cytoplasm, with greater intensity in the nucleus (Fig. S2A). Treatment with DNase I and RNase A confirmed the specificity of BG4 staining (Fig. S2B, C). Gar2 predominantly localized in the nucleolus, consistent with previous reports (Fig. S2D)<sup>23</sup>, and showed strong nuclear colocalization with BG4 signals (Fig. 1A). Quantitative colocalization revealed a Pearson's correlation coefficient of 0.7. Mander's coefficients indicated that 56% of the G4 signal overlapped with Gar2, while 97% of Gar2 signal overlapped with BG4 signal (Fig. 1B). This approach enabled direct visualization of Gar2–G4 colocalization and suggest that Gar2 is predominantly localized to G4-containing regions, but that these are only a subset of all G4s in cells.

To further examine the spatial relationship between Gar2 and G4s *in vivo*, we employed BG4-IEM to achieve nanometer-scale colocalization analysis in *S. pombe*. IEM approaches have previously

been used to detect viral G4s in cells using the 1H6 antibody<sup>24</sup>, a monoclonal antibody that recognizes G4s<sup>25</sup>, supporting the feasibility of ultrastructural G4 visualization. As BG4 had not previously been applied in IEM, we first optimized and validated its suitability (Fig. S3A, B). Omission of BG4 produced no background staining (Fig. S3B), confirming the specificity of the gold-conjugated secondary antibodies. BG4 staining showed nuclear and cytoplasmic localization (Fig. S3A), HA staining was specific to the Gar2–HA strain (Fig. S3C) and absent in the wild-type no epitope-tag control (Fig. S3D). Dual labeling with BG4 and anti-HA antibodies revealed clear nuclear colocalization of Gar2 and G4s (Fig. 1C). Representative IEM images at different magnifications are shown in Fig. S4. Analysis of individual colocalization events demonstrated a mean Gar2–G4 distance of  $11 \pm 3.8$  nm, which is well within the proximity ligation assay detection range ( $\sim 40$  nm)<sup>26</sup> (Fig. 1D). Quantitative analysis of 30 cells from two biological replicates confirmed the reproducibility of Gar2–G4 association, with  $7.91 \pm 0.2\%$  ( $p = 0.0002$ ) of G4s colocalizing with Gar2 and  $32.66 \pm 1.04\%$  ( $p = 0.0009$ ) of Gar2 colocalizing with G4s (Fig. 1E), indicating Gar2 enrichment at a subset of G4-containing sites, consistent with the immunofluorescence analysis (Fig. 1A). Immunofluorescence analysis of cells of varying lengths (a proxy for cell cycle stages) showed Gar2–G4 colocalization across the examined population (Fig. S5). While not based on synchronized cultures, these observations suggest that the interaction is not confined to a single cell cycle stage.

Gar2 has a highly charged N-terminus which likely mediates interactions with nucleolar proteins involved in ribosome biogenesis. Its C-terminus contains two RRM and an RGG motif. Pairwise sequence alignment revealed 26.18% identity between Gar2 and NCL (Fig. S6). Although sequence identity is moderate, domain folding is conserved according to AlphaFold predictions, with RRM domains preserved across these proteins (Fig. S7A). The maximum alignment was observed across RRM1 of Gar2 and RRM3 of NCL (Fig. S7B). Core regions align excellently (RMSD 0.712 Å across 57 pruned Cα pairs),





**Figure 3.** CD spectra of (A) rDNA G4 and (B) c-MYC G4 oligonucleotides titrated with full-length Gar2 protein. (C) CD spectra of c-MYC G4 oligonucleotide with R1R2RGG and R1R2 variants. GraphPad Prism was used to smoothen the curves. The sequences of the oligonucleotides are indicated on the top of each graph.

with blue regions indicating high structural conservation (Fig. S7C). Flexible linkers and missing residues (lime green) account for elevated overall RMSD (55.2 Å). This selective conservation supports shared rRNA binding function.

To validate the direct and specific binding of Gar2 to G4s, we performed *in vitro* assays using purified full-length Gar2 protein (Fig. S8). To examine Gar2–G4 interaction, we performed electrophoretic mobility shift assay (EMSA) with Cy5-labeled DNA oligonucleotides using a previously characterized G4 from the rDNA region of *S. pombe* genome and a mutated counterpart (rDNA M4) as a non-G4 control. Circular dichroism (CD) spectroscopy analysis confirmed that Cy5-rDNA-G4 forms a parallel G4 structure with a positive peak ~260 nm and that Cy5-rDNA-M4 do not form a G4 structure (Fig. 2B). Gar2 binds both substrates in a concentration-dependent manner but binding to rDNA M4 was significantly weaker than rDNA G4 (Fig. 2C), suggesting that Gar2 has higher affinity for the G4 structure.

To assess whether Gar2 recognizes RNA G4s, we tested a well-characterized RNA G4 from the human NRAS gene (NRAS G4) and its mutated non-G4 RNA control (NRAS M4). The NRAS substrate was used to assess structure-dependent rather than sequence-specific recognition by Gar2. CD analysis confirmed a parallel topology for NRAS G4, while the NRAS M4 did not form a G4 structure (Fig. 2D). EMSA demonstrated strong, concentration-dependent Gar2 binding to NRAS G4, while NRAS M4 binding was markedly reduced (Fig. 2E), indicating a clear preference for the structured RNA. Binding affinities were further quantified by microscale thermophoresis (MST). Gar2 bound rDNA G4 and rDNA M4 with  $EC_{50}$  values of  $656 \pm 67$  nM and  $1054 \pm 273$  nM, respectively. For NRAS G4 and NRAS M4,  $EC_{50}$  values were  $162 \pm 21$  nM and  $289 \pm 83$  nM, respectively (Fig. 2F, G). Although the difference in  $EC_{50}$  values between NRAS G4 and NRAS M4 was more modest than suggested by EMSA, both approaches consistently indicate preferential binding to RNA G4 structures. Together, these results suggest that Gar2 binds RNA more strongly than DNA, with an overall preference for RNA G4s over DNA G4s. These findings are in line with previous reports describing NCL's higher specificity for RNA G4s compared to DNA G4s<sup>27</sup>.

To further validate specificity between G4 and non-G4, we performed a competitive EMSA using <sup>32</sup>P-labeled rDNA G4 (1 nM) and 80 nM Gar2 protein. Increasing concentrations of unlabeled rDNA G4 or rDNA M4 were used as competitors. Unlabeled rDNA G4 effectively competed at 10 nM (10-fold excess), whereas rDNA M4 failed to compete even at 250 nM (250-fold excess), confirming

Gar2's specificity for G4 structures (Fig. 2H and I).

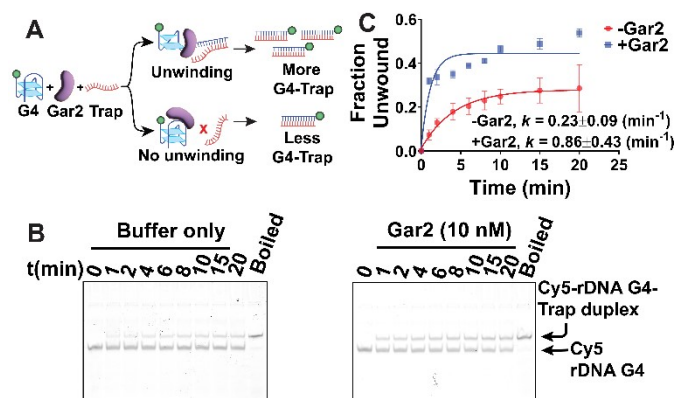
View Article Online

DOI: 10.1039/D6CC01453D

To assess Gar2's specificity for different G4 sequences, and to determine whether its recognition is driven by general structural features of G4 DNA or influenced by sequence context, we selected three well-characterized G4-forming sequences, Pu24T c-MYC, (G4C2)<sub>4</sub>, and 24TTG hTel. CD spectroscopy verified G4 formation under our experimental conditions (Fig. S9A). EMSA demonstrated similar Gar2 binding to Pu24T and (G4C2)<sub>4</sub>, but lower than rDNA G4 (Fig. S9B and Fig. 2C). MST analysis revealed similar binding affinities to the Pu24T c-MYC and the (G4C2)<sub>4</sub>, whereas binding to 24TTG hTel could not be reliably quantified (Fig. S9C). These findings indicate that Gar2's affinity, similar to NCL<sup>28,29</sup>, varies not only between DNA and RNA, but also with different G4 sequences.

To dissect the contribution of individual domains, we cloned and purified two Gar2 truncation mutants, R1R2RGG (Gar2 lacking the N-terminus) and R1R2 (lacking both the N-terminus and RGG motif) (Fig. S8D, E, and Fig. S10A). EMSA with Cy5-rDNA G4 showed that R1R2RGG retained G4 binding, but with reduced efficiency compared to full-length Gar2 (Fig. S10B). Its interaction with the non-G4 rDNA M4 was also weakened, implying that the N-terminus, though intrinsically disordered, enhances nucleic acid binding but is not the primary determinant of G4 specificity (Fig. 2C, S10B, S10C). In contrast, the R1R2 mutant displayed negligible binding to either rDNA G4 or rDNA M4 across the tested concentration range (Fig. S10D and E). Only at very high concentrations a faint shifted band was observed (data not shown), indicating that the RGG domain is critical for efficient binding to both G4 and non-G4 substrates.

Building on known G4-protein interactions in *S. pombe*, where arginines in RGG motifs of Ded1 and Dbp3 are critical for G4 interactions<sup>15</sup>, and in NCL, where phenylalanines are important for G4 interactions<sup>16</sup>, we generated two RGG mutants: R>A mutant (arginines to alanines), and F>A mutant (phenylalanines to alanines) (Fig. S10F; Fig. S8B, C and E). Quantitative EMSA analysis using <sup>32</sup>P-end labeled rDNA G4 showed that wild-type Gar2 binds rDNA G4 with high affinity (apparent  $K_d = 5.7 \pm 2.6$  nM). The F>A mutant displayed only a modest reduction in affinity (apparent  $K_d = 8.1 \pm 2.1$  nM), indicating that, unlike N-terminus truncated NCL<sup>16</sup>, phenylalanines in full-length Gar2 are not major contributors to G4 binding. In contrast, the R>A mutant exhibited weaker interaction with rDNA G4 (apparent  $K_d = 21.3 \pm 10$  nM), indicating the arginine residues in RGG are critical for G4 recognition (Fig. S10G). Together, our results demonstrate that Gar2 requires both its RRM and RGG motif to bind



**Figure 4.** Gar2 unfolds rDNA G4. (A) Schematic illustration of the G4 unwinding trap assay. (B) The G4 unfolding assay was performed in the presence of trap oligo complementary to rDNA G4. Unfolding was analyzed by loading aliquots from different time intervals onto a PAGE gel. (C) The amount of unfolded G4 was quantified from two independent experiments by measuring band intensities using imageJ. The average amount was fitted according to one-phase association function. The  $k$  values represent the mean  $\pm$ SE from two independent experiments.



G4s effectively, with arginines in the RGG playing important roles. Consistently, genetic analyses of the NCL homolog Nsr1 in *Saccharomyces cerevisiae* confirm the necessity of the RGG motif for G4 interaction<sup>30</sup>.

To investigate the functional outcome of Gar2-G4 interaction, we performed CD analysis using rDNA G4 and c-MYC G4 as substrates. In the absence of Gar2, both G4s exhibited characteristic spectra of parallel topology, with a prominent positive peak at ~260 nm. Upon increasing Gar2 concentrations, the amplitude of this peak decreased for both G4s (Fig. 3A, B), indicating structural alterations consistent with G4 binding and suggesting partial G4 destabilization. Furthermore, CD spectra revealed that the R1R2RGG mutant induced similar spectral changes in c-MYC G4 as the full-length Gar2, whereas the R1R2 mutant had no significant effect (Fig. 3C). These findings further confirm the role of the RGG motif in binding the G4.

To directly assess whether Gar2 promotes G4 unfolding, we employed a G4 destabilization assay using Cy5-rDNA G4 as substrate<sup>15, 31</sup>. In this assay, a complementary 'trap' oligonucleotide was included to capture unfolded G4 and prevent G4 refolding, enabling measurement of protein-induced G4 destabilization. In this setup, spontaneous G4 unfolding is typically slow, but can be accelerated by the presence of a G4-destabilizing protein (Fig. 4A). Our data demonstrated that Gar2 increased the rate of G4 unfolding compared to the trap only control (Fig. 4B), with apparent G4 unfolding rate constants (k) of  $0.86 \pm 0.43 \text{ min}^{-1}$  for Gar2 and  $0.23 \pm 0.09 \text{ min}^{-1}$  for trap only (Fig. 4C). These results suggest that Gar2 facilitates G4 destabilization *in vitro*. Notably, Gar2's G4-destabilizing activity stands in contrast to the reported G4-stabilizing function of human NCL<sup>32</sup>, despite both proteins containing conserved RRM and RGG motifs. Recent findings with truncated NCL R1R2R3R4RGG indicate a slight destabilizing effect on the BCL2 G4, which is enhanced in RGG mutants where phenylalanine residues are mutated to alanine<sup>16</sup>, suggesting that NCL family members may exhibit diverse effects on G4s depending on sequence context. In both *S. pombe* and *S. cerevisiae*, G4 motifs are found exclusively on the non-transcribed strand of rDNA<sup>20, 33</sup>. In this context, we speculate that Gar2 contributes to nucleolar RNA regulation and G4 dynamics, thereby supporting ribosome biogenesis in *S. pombe*. By resolving G4s within rDNA and pre-rRNA, Gar2 may facilitate RNA Polymerase I transcription and pre-rRNA processing, potentially contributing to efficient ribosome biogenesis. Notably, this contrasts with the human homolog NCL, which has been reported to bind and stabilize G4 structures, suggesting a potential functional divergence in G4 regulation within the family. Together, these findings support the concept that G4 regulation is both species-specific and context-dependent, with nucleolar proteins adopting distinct functional roles in modulating G4 stability across evolutionary systems.

In conclusion, our study establishes Gar2 as a novel G4-destabilizing protein in fission yeast, revealing a previously unrecognized mechanism by which nucleolar proteins can modulate G4 structures. These findings not only expand our understanding of G4 biology in unicellular eukaryotes but also suggest that the functional diversity of NCL family proteins in G4 regulation may be broader than previously appreciated. Our data further indicate that both RRMs and the RGG motif are critical for efficient G4 recognition, while the intrinsically disordered N-terminal region enhances binding, highlighting the need for structural studies of full-length Gar2 and other NCL homologs to clarify domain-specific contributions. Finally, future analyses of Gar2 will be important for deciphering the molecular logic underlying G4 recognition and remodeling, and for understanding how nucleolar processes are integrated with genome architecture through G4 dynamics.

## Conflicts of interest

There are no conflicts of interest to declare.

## Data availability

The data supporting this article are included in the supplementary information (SI).

## Acknowledgement

We would like to thank Dr. Ikenna Obi and Dr. Pallabi Sengupta for their feedback on this work, and Dr. Tamilselvi Shanmugam for her valuable contributions during the initial stages of this study. We also acknowledge Biochemical Characterization Umeå (BMCU) for CD instrument, Protein Production Sweden at Umeå University (funded by the Swedish Research Council) for BG4 purification and cloning support, the Biochemical Imaging Center (BICU) at Umeå University and the National Microscopy Infrastructure, NMI (VR-RFI 2019-00217) for providing assistance in microscopy, and Umeå Center for Electron Microscopy (UCEM) for IEM support. N.S. received support from Swedish Cancer Society (25 4707 Pj 01 H), Swedish Research Council (VR-MH 2025-02666), and Knut and Alice Wallenberg foundations (KAW 2021.0173).

## Author contributions

Both authors conceived the project and designed the experiments. A.K. performed the experimental studies. N.S. supervised the work and acquired funding for the project. Both authors wrote the manuscript and approved the final version.

## Notes and references

1. Y. Chen and D. Yang, *Curr Protoc Nucleic Acid Chem*, 2012, Chapter 17, Unit17 15.
2. E. Largy, J. L. Mergny and V. Gabelica, *Met Ions Life Sci*, 2016, 16, 203-258.
3. S. Neidle, *Journal of Medicinal Chemistry*, 2016, 59, 5987-6011.
4. B. Zhou, C. Liu, Y. Geng and G. Zhu, *Scientific Reports*, 2015, 5, 16673.
5. N. Kosiol, S. Juranek, P. Brossart, A. Heine and K. Paeschke, *Mol Cancer*, 2021, 20, 40.
6. M. Zhu, J. Gao, X. J. Lin, Y. Y. Gong, Y. C. Qi, Y. L. Ma, Y. X. Song, W. Tan, F. Y. Li, M. Ye, J. Gong, Q. H. Cui, Z. H. Huang, Y. Y. Zhang, X. J. Wang, F. Lan, S. Q. Wang, G. Yuan, Y. Feng and M. Xu, *Nucleic Acids Res*, 2021, 49, 2522-2536.
7. M. Métifiot, S. Amrane, S. Litvak and M.-L. Andreola, *Nucleic Acids Research*, 2014, 42, 12352-12366.
8. D. Rhodes and H. J. Lipps, *Nucleic Acids Res*, 2015, 43, 8627-8637.
9. K. G. Zyner, A. Simeone, S. M. Flynn, C. Doyle, G. Marsico, S. Adhikari, G. Portella, D. Tannahill



- and S. Balasubramanian, *Nat Commun*, 2022, 13, 142.
10. W. F. Yuan, L. Y. Wan, H. Peng, Y. M. Zhong, W. L. Cai, Y. Q. Zhang, W. B. Ai and J. F. Wu, *Cell Biochem Funct*, 2020, 38, 524-532.
11. J. Robinson, F. Raguseo, S. P. Nuccio, D. Liano and M. Di Antonio, *Nucleic Acids Res*, 2021, 49, 8419-8431.
12. V. Gonzalez, K. Guo, L. Hurley and D. Sun, *J Biol Chem*, 2009, 284, 23622-23635.
13. M. Ghosh and M. Singh, *Nucleic Acids Res*, 2018, 46, 10246-10261.
14. V. Meier-Stephenson, *Biophys Rev*, 2022, 14, 635-654.
15. K. K. Yan, I. Obi and N. Sabouri, *Nucleic Acids Res*, 2021, 49, 8339-8354.
16. T. Masuzawa and T. Oyoshi, *ACS Omega*, 2020, 5, 5202-5208.
17. L. Chen, J. Dickerhoff, K. W. Zheng, S. Erramilli, H. Feng, G. Wu, B. Onel, Y. Chen, K. B. Wang, M. Carver, C. Lin, S. Sakai, J. Wan, C. Vinson, L. Hurley, A. A. Kossiakoff, N. Deng, Y. Bai, N. Noinaj and D. Yang, *Science*, 2025, 388, eadr1752.
18. T. Santos, G. F. Salgado, E. J. Cabrita and C. Cruz, *Trends Cell Biol*, 2022, 32, 561-564.
19. G. Biffi, D. Tannahill, J. McCafferty and S. Balasubramanian, *Nat Chem*, 2013, 5, 182-186.
20. N. Sabouri, J. A. Capra and V. A. Zakian, *BMC Biol*, 2014, 12, 101.
21. I. Obi, M. Rentoft, V. Singh, J. Jamroskovic, K. Chand, E. Chorell, F. Westerlund and N. Sabouri, *Nucleic Acids Research*, 2020, 48, 10998-11015.
22. I. Obi, P. Sengupta and N. Sabouri, *Nucleic Acids Res*, 2025, 53.
23. T. Shimada, A. Yamashita and M. Yamamoto, *Mol Biol Cell*, 2003, 14, 2461-2469.
24. S. Artusi, R. Perrone, S. Lago, P. Raffa, E. Di Iorio, G. Palu and S. N. Richter, *Nucleic Acids Res*, 2016, 44, 10343-10353.
25. A. Henderson, Y. Wu, Y. C. Huang, E. A. Chavez, J. Platt, F. B. Johnson, R. M. Brosh, Jr., D. Sen and P. M. Lansdorp, *Nucleic Acids Res*, 2014, 42, 860-869.
26. M. Hegazy, E. Cohen-Barak, J. L. Koetsier, N. A. Najor, C. Arvanitis, E. Sprecher, K. J. Green and L. M. Godsel, *Curr Protoc Cell Biol*, 2020, 89, e115.
27. Y. Khan, T. Azam, J. S. Sundar, S. Maiti and M. K. Ekka, *Biochemistry*, 2023, 62, 1249-1261.
28. S. Lago, E. Tosoni, M. Nadai, M. Palumbo and S. N. Richter, *Biochim Biophys Acta Gen Subj*, 2017, 1861, 1371-1381.
29. A. Saha, P. Duchambon, V. Masson, D. Loew, S. Bombard and M. P. Teulade-Fichou, *Biochemistry*, 2020, 59, 1261-1272.
30. S. Singh, A. Berroyer, M. Kim and N. Kim, *Genetics*, 2020, 216, 1023-1037.
31. J. Gao, A. K. Byrd, B. L. Zybaylov, J. C. Marecki, M. J. Guderyon, A. D. Edwards, S. Chib, K. L. West, Z. J. Waldrip, S. G. Mackintosh, Z. Gao, A. A. Putnam, E. Jankowsky and K. D. Raney, *Chem Commun (Camb)*, 2019, 55, 4467-4470.
32. V. Gonzalez and L. H. Hurley, *Biochemistry*, 2010, 49, 9706-9714.
33. J. A. Capra, K. Paeschke, M. Singh and V. A. Zakian, *PLoS Comput Biol*, 2010, 6, e1000861.



## Data availability statement

The data supporting this article are included in the supplementary information (SI).

